



## Programa BID Caribbean | Proyecto BID-CA2020-041-INS

### TALLER | «Movilización de datos primarios institucionales de biodiversidad venezolana: gestión y publicación».

#### DESCRIPCIÓN

- **Duración:** 2 sesiones de trabajo. Día 1: 2 horas. Día 2: 3 horas.
- **Horario:** **Día 1:** martes 2/11/2021. 9:00 - 11:00 am. **Día 2:** jueves 4/11/2021. 9:00 am - 12:00 m.
- **Modalidad:** virtual a través de la plataforma Zoom: [Taller de capacitación GBIF](#)  
ID de reunión: 824 1703 9410. Código de acceso: 556431.
- **Dirigido a:** responsables de proyectos institucionales que generan sistemáticamente datos primarios sobre biodiversidad venezolana. Se espera colaboración y sinergia entre los participantes.
- **Participantes:** [Alejandro Díaz Petit](#), [Arlene Cardozo Urdaneta](#), [Juan Carlos Amilibia](#), [José Manuel Briceno](#), [jearanguren94@gmail.com](#), [Lila Garcia](#), [Ricmary Bron B](#), [Rodrigo Lazo](#), [Luis Arrieta](#).
- **Facilitadores:** [Haidy Rojas](#), [Alejandro Gallardo](#).

#### JUSTIFICACIÓN DEL TALLER

Durante casi tres décadas, [Provita](#) ha ejecutado cientos de proyectos que van desde la recuperación de especies amenazadas de extinción, hasta la implementación de alternativas de desarrollo sustentable. Nuestra acción institucional ha estado presente en las islas del Caribe, las cordilleras de la Costa y los Andes, y la Amazonía, y se extiende a otras partes del mundo a través de proyectos científicos interinstitucionales, incluyendo la Red Amazónica de Información Socioambiental Georreferenciada ([RAISG](#)) y la Unión Internacional para la Conservación de la Naturaleza ([UICN](#)), principalmente en las comisiones de Supervivencia de Especies ([CSE](#)) y la de Gestión de Ecosistemas ([CGE](#)). Nuestra amplia experiencia en la generación y el uso de diferentes tipos de datos sobre nuestra biodiversidad, junto con una sólida red de aliados y colaboradores institucionales y de la sociedad civil, nos permite establecer acuerdos hacia el intercambio de datos de alta calidad bajo un trabajo colaborativo a nivel país.

#### OBJETIVO DEL TALLER

El objetivo de este taller de trabajo con los líderes de los proyectos coordinados por Provita junto a sus aliados y colaboradores, es fortalecer su base de conocimientos para gestionar la estandarización, sistematización y publicación de los datos primarios sobre la biodiversidad venezolana. En nuestro caso específico, aquellos vinculados a las evaluaciones del estado de conservación de especies y ecosistemas venezolanos, así como los generados institucionalmente, por proyectos de monitorización y conservación de especies.

Con los resultados de este taller se espera iniciar el desarrollo de una política institucional de manejo de datos de biodiversidad que integre las necesidades de los proyectos con los estándares DarwinCore para su publicación a través de los sitios web de Provita y GBIF.org.

## **TALLER | «Movilización de datos primarios institucionales de biodiversidad venezolana: gestión y publicación».**

### **CONTENIDO**

#### **DIA 1**

##### **1. INTRODUCCIÓN AL TALLER**

**Facilitador:** Haidy Rojas.

**Duración estimada:** 15 minutos. Se le solicita a los participantes que previamente vean los videos y preparen sus respuestas a las preguntas / consultas indicadas.

1.1. **Objetivo:** Introducir al participante en el taller.

1.1.1. **Tema:** Introducción al taller.

1.1.1.1. **VIDEO 1:** [Importancia de la publicación de datos de biodiversidad](#) (video 13:02 minutos).

1.1.1.2. **VIDEO 2:** [¿Qué es GBIF?](#) (video 3:00 minutos).

1.1.1.3. [Curso Introducción a GBIF](#) (**opcional**, prerequisite para todos los cursos GBIF).

1.1.1.4. Programa del taller.

1.1.1.5. **Pregunta 1 a los participantes:** [¿Cuál es su expectativa del taller?](#)

##### **2. GESTIÓN DEL PROYECTO**

**Facilitador:** Haidy Rojas.

**Duración estimada:** 45 minutos. Se le solicita a los participantes que previamente vean los videos y preparen sus respuestas a las preguntas / consultas indicadas.

2.1. **Objetivo:** Aprender los conceptos clave (fundamentos) en informática de la biodiversidad, en particular sobre la gestión de datos digitales.

2.1.1. **Tema:** Fundamentos básicos en la informática de la biodiversidad.

2.1.1.1. **VIDEO 3:** [Definiciones](#) (video 12:03 min): datos, campos, etiquetas, listas, carácter.

2.1.1.2. **VIDEO 4:** [Software](#) (video 5:58 min): software a considerar en un proyecto de movilización de datos.

2.1.1.3. **VIDEO 5:** [Estructuras](#) (video 13:10 min): tipos de datos, estructura básica de una base de datos, integridad y seguridad de una base de datos.

2.1.1.4. **VIDEO 6:** [Calidad de los datos](#) (video 12:26 min): aptitud para el uso y medidas de calidad.

2.1.1.5. **VIDEO 7:** [Documentación](#) (video 9:48 minutos): mapeo de los datos, planificación de la movilización de los datos, documentación de la metadata del conjunto de datos.

2.1.1.6. **Pregunta 2 a los participantes:** ¿Tienen alguna duda o comentario respecto a alguno de los conceptos clave relacionados con la informática de la biodiversidad?

2.2. **Objetivo:** Desarrollar una estrategia de movilización de datos adaptada al marco institucional de Provita y a los proyectos de la Línea de Trabajo / Programa.

2.2.1. **Tema:** Política institucional de Provita.

- 2.2.1.1. Misión, visión y valores institucionales.
- 2.2.2. **Tema:** [Consideraciones de Comunicación GBIF](#).
- 2.2.2.1. **Asignación post-taller / consulta a los líderes de proyectos:** para cada conjunto de datos, ¿qué tipo de licencia debe considerarse institucionalmente para la publicación y compartición de la información? Ver [GBIF | Condiciones de Uso](#), esta información será solicitada en la metadata de cada conjunto de datos publicado.
- 2.2.3. **Tema:** Planificación de actividades de la movilización de datos.
- 2.2.3.1. Identificación de actores clave: instituciones y participantes, roles y responsabilidades.
- 2.2.3.2. Flujo de trabajo: planificación, captura de los datos, gestión de los datos y publicación del conjunto de datos. Identificación de cuellos de botella, fortalezas y debilidades. Propuestas.
- 2.2.3.3. **Asignación post-taller / ejercicio práctico:** para cada conjunto de datos, identificar los [actores clave del proyecto](#) de sistematización de la información, sus roles (esta información será indicada en la metadata de cada conjunto de datos publicado) y responsabilidades específicas. Detección de cuellos de botella y vacíos. Propuestas.

### 3. CAPTURA DE DATOS

**Facilitador:** Alejandro Gallardo.

**Duración estimada:** 60 minutos. Se le solicita a los participantes que previamente vean los videos y preparen sus respuestas a las preguntas / consultas indicadas.

- 3.1. **Objetivo:** Identificar los tipos de datos y la mejor manera de capturar la información relevante utilizando las mejores prácticas, software existentes, herramientas y técnicas.
- 3.1.1. **Tema:** Captura de datos.
  - 3.1.1.1. Datos a partir de imágenes y en campo.
  - 3.1.1.2. Problemas comunes de la captura de datos.
  - 3.1.1.3. Introducción al concepto de estandarización.
- 3.1.2. **Tema:** Estandarización de datos. Concepto. Introducción al estándar Darwin Core (DwC) y sus componentes.
  - 3.1.2.1. [Estándares de datos](#).  
[¿Qué es Darwin Core y por qué es importante?](#)  
[List of Darwin Core terms, Darwin Core quick reference guide](#).
- 3.1.3. **Tema:** ¿Qué información se encuentra en GBIF y cómo puedo disponer de ella?
  - 3.1.3.1. [Tipos de conjuntos de datos admitidos](#): metadata, registros de presencia de especies, listas de especies, registros de presencia de especies de un evento de muestreo. Datos primarios según la fuente: especímenes (vivos, preservados o fósiles), literatura, observaciones (humanas o de máquinas).
  - 3.1.3.2. [Cómo hacer búsquedas y descargar datos de biodiversidad desde GBIF](#).
- 3.1.4. **Tema:** Uso de los datos de GBIF.
  - 3.1.4.1. **Asignación práctica 1 para la próxima sesión:** cada participante debe identificar el potencial de la información disponible en GBIF para sus proyectos. Esta información será comentada al resto del grupo en la siguiente sesión del taller.  
Ejemplos relacionados pueden ser indicados como referencia en el [Formulario Datos Biodiversidad Provita](#) (columna O, anaranjada).
  - 3.1.4.2. **Asignación práctica 2 para la próxima sesión:** cada participante debe compartir las [bases de datos que potencialmente se publicarán en GBIF](#), con todos o al menos 10 registros, que permitan describir el contenido de cada campo de

información. La primera fila tendrá el término propuesto DwC y la segunda fila el término original.

El hipervínculo a cada conjunto de datos debe estar referido en el [Formulario Datos Biodiversidad Provita](#) (columna P, verde).

## DÍA 2

### 4. CONTINUACIÓN: CAPTURA DE DATOS

**Facilitador:** Alejandro Gallardo.

**Duración estimada:** 45 minutos.

- 4.1.1. **Tema:** sesión de preguntas y respuestas para aclarar cualquier duda sobre el material del día anterior.
  - 4.1.1.1. **Revisión de asignación práctica 1:** [Formulario Datos Biodiversidad Provita](#) (columna O, anaranjada).
  - 4.1.1.2. **Revisión de asignación práctica 2:** [Formulario Datos Biodiversidad Provita](#) (columna P, verde).

### 5. GESTIÓN DE DATOS

**Facilitador:** Alejandro Gallardo.

**Duración estimada:** 45 minutos.

- 5.1. **Objetivo:** Aprender conceptos relacionados con la calidad de los datos y recibir una introducción a las herramientas utilizadas para estandarizar, validar, limpiar y mejorar los datos.
  - 5.1.1. **Tema:** [Requisitos de calidad de los datos](#).
    - 5.1.1.1. Campos obligatorios y opcionales por cada tipo de conjunto de datos.
    - 5.1.1.2. Metadata asociada al conjunto de datos.
  - 5.1.2. **Tema:** Herramientas de software diseñadas para la limpieza de datos de biodiversidad.
    - 5.1.2.1. [OpenRefine](#), **VIDEO 8** (opcional): [Depuración de datos con Open Refine](#) (video 51:34 minutos).
    - 5.1.2.2. Microsoft Excel.
    - 5.1.2.3. **VIDEO 9** (opcional): [Documentación y herramientas de georreferenciación](#) (video 1:35 horas).
    - 5.1.2.4. [Aplicaciones de Google](#): Google Map, Google Earth, Hoja de Cálculo, Drive, entre otros.
  - 5.1.3. **Tema:** Herramientas de software diseñadas para validar y mejorar los datos de biodiversidad.
    - 5.1.3.1. Herramientas disponibles en [GBIF](#)
    - 5.1.3.2. [Biodiversity Information Standards \(TDWG\)](#)
    - 5.1.3.3. [Global Biodiversity Information Facility · GitHub](#)
  - 5.1.4. **Tema:** Referencias que validan y mejoran los datos de biodiversidad.
    - 5.1.4.1. Taxonomía de las especies:
      - 5.1.4.1.1. taxonID: <https://www.gbif.org>
      - 5.1.4.1.2. identificationReferences: [iNaturalist](#) | [Flickr](#)
      - 5.1.4.1.3. kingdom, phylum, class, order, family, genus, specificEpithet, scientificName, AuthorshipscientificNameID, nomenclaturalCode, taxonomicStatus, nomenclaturalStatus: [Plants of the World Online](#) | [Kew Science](#) | [International Plant Names Index \(IPNI\)](#) | [Integrated Taxonomic Information System](#) | [Amphibian Species of the World](#) | [The Code Online](#) | [IAPT](#) | [The Code](#)

- 5.1.4.2. Geographic information: higherGeographyID, higherGeography, continent, waterbody, islandGroup, Island, country, countryCode, stateProvince, county, locationAccordingTo:
  - 5.1.4.2.1. Controlled vocabulary: the Getty Thesaurus of Geographic Names: [Getty Thesaurus of Geographic Names® Online](#)
  - 5.1.4.2.2. Otros

**DESCANSO:** 15 minutos

## 6. PUBLICACIÓN DE DATOS

**Facilitador:** Haidy Rojas.

**Duración estimada:** 45 minutos.

- 6.1. **Objetivo:** Conocer el proceso de facilitar que los datos de biodiversidad estén disponibles en línea libremente, también conocido como publicación de datos, utilizando la herramienta de publicación integrada de GBIF, IPT (*Integrated Publishing Toolkit*).
  - 6.1.1. **Tema:** Definir los tipos y subtipos de datos publicables (si los hay) para un conjunto de datos de biodiversidad.
    - 6.1.1.1. [¿Qué es un Archivo Darwin Core \(DwC-A\) y por qué es importante?](#)
  - 6.1.2. **Tema:** ¿Qué puedo publicar como Provita en GBIF?
    - 6.1.2.1. Resultados del diagnóstico de la consulta a los líderes de proyectos institucionales, en octubre-noviembre de 2021: [conjunto de datos potencialmente publicables en GBIF](#).
    - 6.1.2.2. Sugerencias a cada Línea de Trabajo.
  - 6.1.3. **Tema:** ¿Dónde publico los datos?
    - 6.1.3.1. Publicaciones institucionales
      - 6.1.3.1.1. Página web institucional
        - 6.1.3.1.1.1. [Provita](#)
        - 6.1.3.1.1.2. [GeoPortal Provita](#)
        - 6.1.3.1.1.3. [Libro Rojo de la Fauna Venezolana](#)
      - 6.1.3.2. Sistema de Información País
      - 6.1.3.3. Nodo País
        - 6.1.3.3.1. MINEC: Sistema Venezolano de Información sobre Diversidad Biológica (SVIDB)
      - 6.1.3.4. Nodo temático
        - 6.1.3.4.1.1. [OBIS](#)
        - 6.1.3.4.1.2. [Andean Network of BioNET-INTERNATIONAL](#)
      - 6.1.3.5. [GBIF](#)
        - 6.1.3.5.1.1. [Guía para la publicación de juegos de datos en GBIF.org](#)
        - 6.1.3.5.1.2. [IPT: Herramientas para la publicación integrada de datos.](#)
          - 6.1.3.5.1.2.1. **IPT:** vista general del recurso. Cargue o conecte los datos de origen, mapee los datos a Darwin Core, complete los metadatos, publique el recurso y regístrelo en la red de GBIF.
          - 6.1.3.5.1.2.2. **VIDEO 10** (opcional): [introducción a la herramienta de publicación IPT \(Integrated Publishing Toolkit\)](#) (video 13:05 minutos).
    - 6.1.3.6. Publicaciones personales
      - 6.1.3.6.1. [iNaturalist](#)
      - 6.1.3.6.2. [eBird](#)

## **OTROS**

[Curso de Movilización de Datos sobre Biodiversidad: glosario](#)

[Curso de Movilización de Datos sobre Biodiversidad: documentación clave](#)

**Pregunta 3 a los participantes:** Considerando su expectativa inicial a este taller, [¿qué propone para mejorarlo?](#) Considere que ahora deberá replicarse a nuestros aliados y colaboradores institucionales.

**Actividad de cierre del taller:** captura de pantalla de todos los participantes en el taller.

## ANEXO

### GESTIÓN DE PROYECTOS

#### PARTES ASOCIADAS AL PROYECTO

Actores clave que participan en el proyecto de movilización de datos y algunos de estos, deberán tener su reconocimiento en la publicación final. Esta información con su detalle, comprende parte de la metadata del conjunto de datos a publicar.

ROLE	CONCEPT
<b>Author</b> Autor	an agent associated with authoring a publication that used the data set, or of a data paper.
<b>Content Provider</b>	an agent who contributed content to a data set (data set being described may be a composite).
<b>Custodian Steward</b>	an agent who is responsible for/takes care of the data set.
<b>Distributor</b>	an agent involved in the publishing/distribution chain of a data set.
<b>Editor</b> Editor	an agent associated with editing a publication that used the data set, or of a data paper.
<b>Metadata Provider</b>	an agent responsible for providing the metadata.
<b>Originator</b>	an agent who originally gathered/prepared the data set.
<b>Owner</b> Propietario	an agent who owns the data set (may or may not be the custodian).
<b>Point Of Contact</b> Contacto	an agent to contact for further information about the data set.
<b>Principal Investigator</b> Investigador Principal Líder del Proyecto	a primary scientific contact associated with the data set.
<b>Processor</b> Analista de datos	an agent responsible for any post-collection processing of the data set.
<b>Publisher</b> Publicador	the agent associated with publishing a publication that used the data set, or of a data paper.
<b>User</b> Usuario	an agent that makes use of the dataset.
<b>Programmer</b> Programador	an agent providing informatics/programming support related to the data set.
<b>Curator</b> Curador	an agent that maintains and documents the specimens in a collection. Some of their duties include preparing and labelling specimens so they are ready for identification and protecting the specimens.

<b>First Name</b>	José	<b>Last Name</b>	Pérez
<b>Position</b>	Líder del Proyecto	<b>Organization</b>	Provita
<b>Address</b>	Calle La Joya entre Av. Francisco de Miranda y Av. Libertador. Edif. Unidad Técnica del Este, Piso 10, Ofic. 29 y 30. AP 47552	<b>City</b>	Caracas
<b>State/Province</b>	Miranda	<b>Country</b>	Venezuela
<b>Postal Code</b>	1060	<b>Phone</b>	+582122637240
<b>Email</b>	jperez@provitaonline.org.ve	<b>Home Page</b>	<a href="https://www.provita.org.ve/">https://www.provita.org.ve/</a>
<b>Personnel Directory</b>	<a href="https://www.researchgate.net/profile/XXXXXXXX">https://www.researchgate.net/profile/XXXXXXXX</a>	<b>Personnel Identifier</b> ID ORCID	<a href="https://info.orcid.org/XXX-X-XXXX-XXXX-XXXX">https://info.orcid.org/XXX-X-XXXX-XXXX-XXXX</a>
<b>Roles</b>	Investigador principal, autor,		
<b>Responsabilidades</b>	Coordinación del proyecto, publicación de la información, punto de contacto		

<b>First Name</b>		<b>Last Name</b>	
<b>Position</b>		<b>Organization</b>	Provita
<b>Address</b>	Calle La Joya entre Av. Francisco de Miranda y Av. Libertador. Edif. Unidad Técnica del Este, Piso 10, Ofic. 29 y 30. AP 47552	<b>City</b>	Caracas
<b>State/Province</b>	Miranda	<b>Country</b>	Venezuela
<b>Postal Code</b>	1060	<b>Phone</b>	+582122637240



<b>Email</b>		<b>Home Page</b>	<a href="https://www.provita.org.ve/">https://www.provita.org.ve /</a>
<b>Personnel Directory</b>		<b>Personnel Identifier</b>	
<b>Roles</b>	Propietaria	<b>Responsabilidades</b>	

### Ejemplos:

Caribbean OBIS Node: Pez León en Venezuela <https://doi.org/10.15468/bi3ixg> accessed via GBIF.org on 2021-10-23.

Gaiani M (2019). Insect Occurrence Data from MIZA. Version 1.1. Universidad Central de Venezuela. Occurrence dataset <https://doi.org/10.15468/ah3is7> accessed via GBIF.org on 2021-10-23.